Data Assimilation Office (DAO) Algorithm Theoretical Basis Document (ATBD)

Revisions and Addenda

Richard B. Rood, Principal Investigator
Robert M. Atlas, Head, Data Assimilation Office

January 20, 2000

A. Introduction and Overview

This document stands as both revisions and addenda to the Algorithm Theoretical Basis Document (ATBD) generated by the Data Assimilation Office in 1996 (DAO, 1996; hereafter ATBD-96). The document is a requirement of Earth Observing System (EOS) Project Office. The plans and algorithms proposed by the DAO in ATBD-96 have to a large extent been followed and implemented with the recognition that the scope of proposed implementations is less than described in ATBD-96.

ATBD-96 can be found at

http://dao.gsfc.nasa.gov/

then "Publications," followed by "Algorithm Theoretical Basis Document."

The current document uses ATBD-96 as a baseline. Details of important specifics of the DAO activities in the year 2000 and beyond are given. There are four major Elements in the current document, which are easily recognizable by the information at the bottom of the pages in the primary documents.

Element 1) Introduction and Overview (Part A). This part of the document provides an overview of the rest of the document and covers important issues that are not represented in the more scientific content of the other Elements. Appendices to this Element include:

- Response to the reviews of ATBD-96 (App A)
  - Fifth Report of the DAO Advisory Panel
- Plan for Joint Activities with NCEP on Data Usage (App B)
- White Paper on Joint Modeling with NCAR (App C)
- Management Plan of Joint NCAR-DAO Proposal (App D)
- Minutes of multi-institutional software management meeting (App E)

Element 2) Description of important subsystems of the Goddard Earth Observing System (GEOS) Data Assimilation System (DAS) that have been implemented in support of the Terra launch (Part B). Significant attention is given to documentation of the improvement of geophysical parameters over earlier versions of the DAS.

Element 3) Key efforts in investigation of new data types (Part C). The ozone data assimilation system, which provides a standard product as of the Terra launch in December 1999, is described here.

Element 4) Description of the algorithms in the prototype general circulation model that has been developed jointly with the National Center of Atmospheric Research (NCAR) (Part D). This section includes representative results from the prototype model as well as a description of the proposed next-generation assimilation system that uses this model.

These four major sections will rely, when possible, on peer-reviewed papers. That is, if there are papers already appearing in the literature, rather than repeating those developments here, only the highlights are given and the peer-reviewed paper serves as the foundation document. In the case of the ozone system, a primary document is a paper recently submitted for peer review, which is included. Other papers, which have yet to appear, are available on request (as are higher definition figures). Largely, the sections associated with Element 2, Element 3, and Element 4 can be reviewed independently of each other

In addition to the integrating and summary discussion given below, also attached as Appendix A to this section is an earlier written response to the reviews provided for ATBD-96. That response addresses all of the major issues raised in ATBD-96, with the explicit recognition that resolution of several of those issues lie outside of the control of the DAO. Scientifically, the reviews of ATBD-96 raised the greatest concerns about the general circulation model used by the DAO. These concerns became part of the motivation to initiate the new model development in partnership with NCAR. Reading the response to the reviews of ATBD-96 will provide continuity from ATBD-96 to the current document. This will be especially relevant to individuals who have been involved in continuing reviews of the DAO.

The GEOS DAS (Goddard Earth Observing System Data Assimilation System) has been developed over a number of years by the DAO with the primary milestones driven by major launches of the EOS series of satellites. Both in support of pre-launch science and to provide incremental foci and evaluations, several versions of the GEOS system have been developed. GEOS-1 was the first multi-year reanalysis produced by any organization and became the first baseline on which to measure improvement. Subsequent reanalyses by the National Centers of Environmental Prediction (NCEP) and the European Center for Medium-range Weather Forecasts (ECMWF) were significant improvements over the GEOS-1 product, lessening the interest in GEOS-1. More than 100 papers and presentations came from GEOS-1, and the GEOS-1 products are still used by a number of researchers, especially in tropospheric and stratospheric chemistry applications.

ATBD-96 started from the GEOS-1 baseline and described projected GEOS-2 and GEOS-3 systems. GEOS-2 was described with significant specificity, and GEOS-3 was described as a series of research developments that might contribute to future versions. At that writing, GEOS-2 was projected as the Terra system, with Terra launch then scheduled for 1997. Since ATBD-96, the version numbers associated with particular launches have changed. In the present numbering system, the configuration of GEOS-3 with 1-degree horizontal resolution is defined as the Terra system and will usually be referred to as GEOS-Terra. Because plans have changed significantly with the projected successes of the new general circulation model, the details of GEOS-3, as defined in ATBD-96, are no longer accurate.

GEOS-Terra (i.e. GEOS-3 in the current version numbering) is quite consistent with the plans and algorithms detailed in Chapter 5 of ATBD-96, and Chapter 5 should be the primary reference document for GEOS-Terra. In particular, the Physical-space Statistical Analysis System (PSAS) has been implemented as described in Chapter 5. While there have been significant improvements in the treatment of error statistics and data preprocessing, many of the more sophisticated developments presented in ATBD-96 have not been implemented. Because of the necessity to assure a DAO product for Terra, as well as to support the Tropical Rainfall and Measurement Mission (TRMM), it was not possible to do a complete change of the general circulation model. Therefore, GEOS-Terra uses a general circulation model, which is an incremental development of the model presented in ATBD-96. Indeed, much of the improvement in the representation of geophysical parameters measured in GEOS-Terra can be attributed to model improvements. This includes the successful implementation of the land-surface model as described in ATBD-96. Because of the broad consistency with the plans presented in ATBD-96, this document (see Element 2, above) includes only a description of the sub-systems actually implemented. The improvements in the product are documented by comparison with older GEOS baselines and other independent data sets.

The documentation of improved performance of GEOS-3 will rely, at times, on various incremental configurations of the GEOS DAS. There have been numerous implementations of GEOS-1 and GEOS-2 in support of specific NASA projects. As in the past, the DAO has supported, with forecast and analysis products, most of the NASA-sponsored aircraft missions to study stratospheric chemistry, as well as some of the tropospheric chemistry missions. GEOS products and algorithms continue to be used in the chemical assessment activities carried out by the Global Modeling Initiative (GMI) at Lawrence Livermore National Laboratory. Given the focus of the Terra mission, however, the most important transitional configuration of GEOS is a GEOS-2 system, GEOS-TRMM. This system was configured in support of the CERES (Clouds and the Earth's Radiant Energy System) instrument on the TRMM satellite. Challenges provided by CERES support led to important improvements in the GEOS DAS that benefit the Terra system. The responses by the DAO to the problems raised by the CERES team will also be detailed in Element 2 (Part B).

An important part of the DAO mission, as well as a crucial path for improving the assimilated data product, is the use of new types of observations. The GEOS-Terra system uses significantly more satellite data than GEOS-1, and these are included in Element 2 (Part B). Several other data types, addressed in Element 3 (Part C), have been addressed in either research versions of the DAS or in off-line versions of the DAS. Off-line assimilation systems use the output from the meteorological assimilation in a sequential matter to assimilate another parameter. The ozone product for Terra is an off-line product. In addition, discussions will be provided on use of observations from the operational infrared and microwave sounders, TRMM, the global positioning system, and scatterometry. Research activities to improve the representation of near-surface and hydrological processes over land will also be described.

The most severe comments about DAO algorithms described in ATBD-96 were directed at the assimilating general circulation model used in GEOS-1 and, then, projected for use in all future versions of the GEOS system. There have been long-standing issues in the development of atmospheric models at Goddard, and external customer expectations place high pressure on both the forecast and climate performance of the GEOS model. This ultimately feeds back to the quality of the underlying physics represented in the model. To address these issues, the DAO and NCAR initiated a joint project to develop a next-generation atmospheric model. The joint project with NCAR was started after numerous discussions with many other government institutions and was based on both shared scientific interest and potential for collaborative success. Because of the significant changes in the underlying structure and formulation of the next-generation model, changes in the basic design of the model-analysis (i.e. PSAS) interface are required. The documents contained in Element 4 (Part D) discuss all aspects of the joint model as well as providing an overview of the design of the new assimilation system. The new design allows us to address some embedded artifacts in the existing design that hinder optimal performance of the GEOS DAS. We anticipate having results from a prototype of the next-generation assimilation system by the time of the panel review.

The remainder of this overview document will discuss important issues that are not addressed in documents Element 2, Element 3, or Element 4. First, joint activities with NCEP on the utilization of both historical and new types of observations will be presented. Then some of the issues of software management and computational resources will be discussed more fully. (See also, Appendix A)

Both Goddard and DAO management are committed to increased use of partnerships to achieve the ambitious goals of the Data Assimilation Office. Due to both personnel and computational resource stresses, there is increased motivation for laboratories sponsored by different agencies to combine resources to work towards goals that share common success criteria. The growing collaboration on model development with NCAR and a consortium of Department of Energy laboratories is an example of this new generation of collaborations. A natural link also exists to many activities within the National Oceanic and Atmospheric Administration (NOAA). The most approachable link with ongoing NOAA activities exists in the improved use of observations. Therefore, scientists from NOAA's National Centers of Environmental Prediction (NCEP), National Environmental Satellite, Data, Information Service (NESDIS), and the DAO have developed a plan to address the issues of data usage with a more unified front.

The details of this plan are given in the document attached as Appendix B of this overview document. The ability to reach common ground with NCEP on use of new data types is grounded on the fact that much of the foundation to address the utility of a particular data type is independent of the details of how a particular organization approaches the modeling and assimilation problem. Therefore, activities in one organization are clearly relevant to activities in the other organization, even if the specific goals of the application are substantively different. For instance, forward models suitable for assimilation applications can be exchanged with relatively little interaction with the rest of the DAS. Furthermore, in all assimilation applications there is strong component of weather forecasting, and, with sufficient planning, forecast impact information from the DAO's DAS is meaningful to scientists at NCEP and vice versa. The potential of increased interaction with NCEP on the development of analysis and modeling systems is great, and DAO and Goddard management are specifically pursuing these links with NCEP and other NOAA-funded activities. This includes more integration of NOAA interests in the joint modeling activities with NCAR.

At the basis of successful joint activities are a number of difficult organizational and management issues that are being addressed more actively than before. Prior to the initiation of the joint activity with NCAR, members of the two organizations tried to identify the organizational and motivational issues that must be addressed for successful collaboration. These are discussed in Appendix C of this overview document. Aside from clarity of purpose and agreements of how organizations will work together, more formal methods of managing and developing scientific software are essential.

One tangible part of Earth-science modeling and assimilation is software. Since complex software systems are needed to represent the interface between scientific principles and the computational environment, a more controlled approach to software development will allow more scientists, more institutions, to work more effectively on common problems. In this way, those things that are common and non-controversial, but that consume significant resources to build and maintain, can be standardized and shared. However, they have to be designed to not only function robustly in increasingly fragile computational environments, but to allow effective scientific investigation and development of algorithms to address those questions that are still under active investigation. Application of the principles of software engineering to scientific development, and especially multi-institutional scientific development, is a difficult and controversial topic. Our commitment to address these problems is expressed in Appendices D and E.

The computational issues facing the DAO are shared by all of the United State's Earth-science activities that require the use of high-performance computers to deliver assimilated data products, forecasts, and climate simulations to customers on a routine basis. There are a number of ways to look at the problem, but there are two overarching issues. The first is that there are tremendous challenges in moving away from Cray (C90)-like shared-memory, vector computers to distributed-memory, potentially massively-parallel computers that use non-vector commodity-based microprocessors. The second is that outside of the United States scientists are not having to face these software issues because they are not reliant on distributed-memory, massively-parallel computers. Therefore, U. S. organizations are at a clear disadvantage in their ability to utilize data and provide simulations, and they have to spend significant portions of their resources in tremendously difficult, high-risk software development. There is both implicit and explicit demand that U. S. organizations provide products of comparable quality.

Figure 1 exemplifies the computational problem faced by U. S. scientists. This figure shows the throughput of only the forecast model of the ECMWF forecast-assimilation system. The VPP and SX values are from Japanese-made supercomputers that are subject to stringent U. S. import limitations, which U. S. scientists, fundamentally, do not have access to. There are three important points.

Point-1) On just the forecast model part of the problem, comparable throughput numbers can be reached on the T3E, available to U. S. scientists, as can be reached on the VPP and the SX.

Point-2) The throughput numbers on the VPP and SX are reached on a relatively small number of processors, 16-24, while the T3E numbers are reached on many hundreds or, perhaps, more than 1000 processors.

Point-3) While the SP curve is not a good indicator of ultimate SP performance, it does reflect the difficulty in obtaining high performance computing from distributed-memory, multiple-processor computers.

There are a number of important other considerations embedded in this plot. The software effort required to achieve these throughout levels is large and performance varies widely from platform to platform. Without devaluing the software quality needed to achieve these throughput levels on a 24 processor VPP, the effort to accomplish the same job on a 1000 processor T3E is much, much more difficult. Furthermore, these numbers are for only the forecast model, and not the assimilation system as a whole, functioning in a product-generation environment. It has proven much more difficult to develop software that will allow the assimilation system as a whole to scale to, for instance, even 64 processors. There are also fundamental issues of computational science (i.e. Amdahl's Law), which suggest that scaling beyond, on the order of, 100 processors is not possible. Therefore, U. S. institutions, the DAO included, are faced with tremendous efforts in software development that may or may not pay off. This effort is spent in the context of the following two facts. It is difficult to recruit and retain computational scientists with the correct expertise. From the point of view of business competitiveness, it is much simpler for non-U. S. organizations to increase their scaling to, for instance, 32 processors, than for the U. S. organizations to add several hundred more processors to achieve the same rate of throughput.

The DAO was out in front recognizing these computational issues and designed their computing facilities to take into account both the availability of hardware platforms and the inability of DAO software to scale to many hundreds of processors. The DAO's main compute engines are a number of 64 processor Origin 2K's. These machines are each basically tasked to a specific suite of problems. With the mandate to provide our Terra-support analyses at the one degree horizontal resolution, our throughput falls to less that five days per day for the data assimilation system. This represents a major fragility in the provision of the DAO products. Perhaps even more important in the long run, validation and testing is severely compromised because external delivery schedules overburden the computational resources. This leads to further compromises in research activities.

The solution to these problems are no easy and straightforward. The DAO takes advantage of what parallelism it can exploit, for instance, running multiple production streams at reduced resolution. Testing is often done at lower resolution with the hope that low resolution results basically transfer to higher resolution. Software improvement activities are underway, but it is difficult to maintain continuous improvement with the various other research and product pressures. The next-generation system brings forward new algorithms that might provide order of magnitude performance improvement. All of these efforts consume resources and are high risk.

The DAO's experience facing these computational issues has been recognized externally. Richard Rood, no longer the Head of the DAO, is assigned to a number of national activities to address these issues from a strategic point of view, and is a recognized leader on modeling-assimilation and its relationship to supercomputing. He spends much of his time on this problem. Robert Atlas runs the DAO and assumes responsibility for the DAO products and the organization's scientific direction.

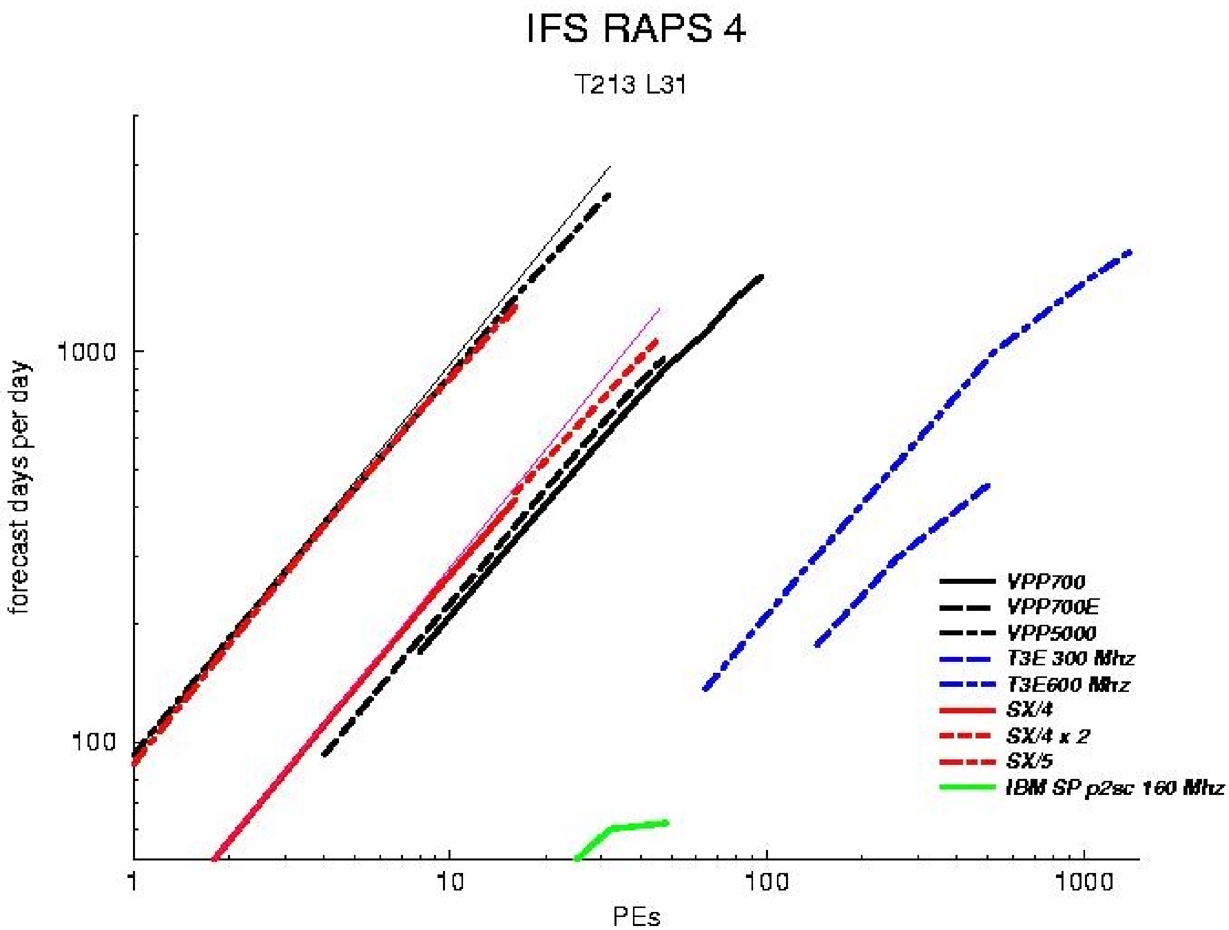

Figure 1: Throughput of the Integrated Forecast System (IFS) of the European Center for Medium-range Weather Forecasts (ECMWF). On the horizontal axis is plotted the number of processor elements. On the vertical axis is throughput measured in number of days forecasted for each day of processor time for the prescribed number of processor elements. The SX machines are from NEC. The VPP machines are from Fujitsu. The T3E from SGI/Cray, and the SP from IBM. The SP numbers were from November 1998, and provided by IBM, but should not be viewed as the representation of current capabilities. These numbers are for the forecast model only and do not include any assimilation cycles. Figure provided by ECMWF.